A Probabilistic Study of Bound Consistency for the Alldifferent Constraint

Jérémie Du Boisberranger[1], Danièle Gardy[1],
Xavier Lorca[2], and Charlotte Truchet[3]

[1] Université de Versailles,PRiSM UMR 8144,
FR-78035 Versailles, France
{Daniele.Gardy,Jeremie.Du-Boisberranger}@prism.uvsq.fr
[2] École des Mines de Nantes, LINA UMR CNRS 6241,
FR-44307 Nantes Cedex 3, France
Xavier.Lorca@mines-nantes.fr
[3] Université de Nantes, LINA UMR CNRS 6241,
FR-44322 Nantes, France
Charlotte.Truchet@univ-nantes.fr

Abstract. This paper introduces a mathematical model for bound consistency of the constraint `AllDifferent`. It allows us to compute the probability that the filtering algorithm effectively removes at least one value in the variable domains. A complete study of the bound consistency properties is then proposed. It identifies several behaviors depending on some macroscopic quantities related to the variables and the domains. Finally, it is shown that the probability for an `AllDifferent` constraint to be bound consistent can be asymptotically estimated in constant time. The experiments illustrate that the precision is good enough for a practical use in constraint programming.

1 Introduction

Constraint Programming (CP) aims at solving hard combinatorial problems expressed as Constraint Satisfaction Problems (CSP): variables are the unknown of the problem, each variable has a domain of possible values and the constraints are logical predicates that link the variables. The CP operational nature is based on the propagation-search paradigm. The search part consists in an exploration of the possible assignments of domains values to the variables, while the filtering/propagation mechanism detects and suppresses inconsistent parts of the domains (*i.e.* values that cannot appear in a solution of the constraints).

In particular, filtering algorithms for global constraints [2], which define logical relations on a subset of the variables, have been intensively studied. They are usually more efficient than the propagation of equivalent expressions as several binary or small arity constraints. One of the major issues in studying propagation is the pursuit of a fair balance between efficiency (time complexity) and effective performance, which can be expressed for instance as the number of inconsistent values detected by the filtering algorithms. It may happen that propagation algorithms are useless (no inconsistent values detected), and, even if they are often efficient, they are called each time a value is assigned to a variable. In the worst case, this makes a number of calls exponential in the number of variables.

This paper introduces a probabilistic model for bound consistency (BC) of the `alldifferent` constraint, relying on the inner combinatorics of this constraint. In this model, it defines the probability that the BC algorithm removes at least one value in the variable domains. Then, a complete study of this model is developed for two different probability distributions for the domains. The interest of this study is twofold: first, the probability computation leads to the identification of several distinct behaviors, depending on some macroscopic quantities of the problem. Secondly, we show that the probability that a BC algorithm is effective can be estimated in constant time. Experiments confirm that this asymptotical approximation is precise enough for a practical use.

This work strongly relies on the combinatorics embedded in many global constraints (*e.g.* at least all the cardinality constraints). This link has also been exploited by [8] to count the number of solutions of a constraint, in order to predict promising areas of the search space and guide search heuristics. Although their goal is different from ours, they also exploit the mathematics and combinatorics inside these constraints. In another way, triggering the filtering algorithm of a global constraint has also been proposed in [4] for the global cardinality constraint. The author determines a threshold, based on the possible variable-values assignments, to determine if the filtering algorithm must be executed. Such an approach leads to a good tradeoff between calculation time and effective filtering.

In the following, Section 2 introduces a probabilistic model of the bound consistency `AllDifferent` constraint. Section 3 presents the mathematical results of the paper, for two distributions of the variables domains. Section 4 criticizes these results, and provides a practical validation. Section 5 concludes this paper and discusses some further works.

2 Probabilistic Approach of Bound Consistency

After some preliminaries definitions and propositions on BC for `alldifferent` and on probabilities, we introduce a probabilistic model for BC of `alldifferent`. Then we define the probability that an `alldifferent` constraint remains BC after an instantiation and give a general formula for it.

Consider an `alldifferent` constraint on n variables $V_1...V_n$, with domains $D_1...D_n$. The size of domain D_i is $d_i = |D_i|$. Let E be the union of all the domains: $E = \cup_{1\leq i\leq n} D_i$, and m its size. For an integer interval E and a set I, we write $I \Subset E$ as a shortcut for : $I \subset E$ and I is an integer interval.

2.1 Bound consistency for `alldifferent`

The `alldifferent` constraint is a well known global constraint for which many consistency algorithms have been proposed. The reader can be refered to Van Hoeve's survey [7] for a state of the art. We focus here on bound consistency, as proposed by [5], and thus assume that all the domains D_i are integer intervals. E can be assumed to be an integer interval without loss of generality (if E is not an interval, it can be cut into different pieces, with independent `alldifferent` constraints for each of them).

Definition 1. *Let $I \Subset E$. Let K_I be the set of variables for which the domains are subintervals of I, $K_I = \{i, D_i \Subset I\}$. I is a* Hall interval *if and only if $|K_I| = |I|$.*

Hall intervals are the subintervals of E that contain just enough values to ensure that every variable of K_I can be assigned a value. Consequently, the other variables (those not in K_I) cannot take their values in I.

Proposition 1. *A* `alldifferent` *constraint on a set of variables* $V_1...V_n$*, a set of domains* $D_1...D_n$ *is* bound-consistent *if and only if the two following conditions are true:*

1. *for* $I \Subset E$, $|K_I| \leq |I|$,
2. *and for* $I \Subset E$, $|K_I| = |I|$ *implies* $\forall i \notin K_I, I \cap D_i = \emptyset$.

In the following, we are interested in what happens during an instantiation. It can be assumed without loss of generality that the instantiation is done for the variable V_n (rename the variables if necessary). V_n is assigned a value $x \in D_n$. We also assume that the binary constraints $\neq$ have been filtered (that is, only the x value has been removed from the other domains, which is equivalent to a Forward Checking on the $\neq$ clique).

The number of variables n is computed dynamically, that is, n is the number of *free* variables before the instantiation. Let E° be the state of E before the instantiation ($E^\circ = D_1 \cup ... \cup D_n$), and $E^\otimes$ be the state of E after the instantiation ($E^\otimes = D_1 \cup ... \cup D_{n-1}$). In the same way, for an interval $I \Subset E$, $I^\circ = I^\otimes \cup \{x\}$ if $I^\otimes \neq I^\circ$ and $I^\otimes$ otherwise.

The following Proposition details in which case the constraint, being BC on $V_1...V_n$, *remains* BC after the instantiation of V_n.

Proposition 2. *Consider a* `alldifferent` *constraint on variables* $V_1...V_n$*, with domains* $D_1...D_n$*, such that the constraint is BC. Assume that the variable* V_n *is set to a value* $x \in D_n$*, and let* v *be the position of* x *within* E° ($v = x - min(E^\circ)$)*. The constraint remains BC if and only if there exists* $I^\otimes \Subset E^\otimes$ *such that* $I^\otimes \neq I^\circ$ *and the two following conditions hold:*

(1) $D_n \subsetneq I^\circ$,

(2) there are exactly $|I^\circ| - 1$ *domains included in* I° *and at least one other domain is neither included in, nor outside* I°.

Proof. The proof is based on the fact that an instantiation cannot create new intersections on the domains. Suppose that the problem is BC, then $\forall I^\circ \Subset E^\circ$, $|K_{I^\circ}| \leq |I^\circ|$. Let $I^\otimes \Subset E^\otimes$, and $l = |I^\otimes|$ and $\alpha = min(I^\otimes) - min(E^\otimes)$.
If $\alpha + l < v$, that is, $I^\otimes = I^\circ$, we have $|K_{I^\otimes}| = |K_{I^\circ}| \leq |I^\circ|$. Indeed, even a domain which had x for right bound before the instantiation, has its right bound out of $I^\otimes$ after. Then, item 2. of Proposition 1 holds by the first remark. In that case, no inconsistency is created. If $\alpha > v$, the argumentation is the same, interchanging left by right. If $\alpha + l \geq v$ and $\alpha \leq v$, that is $I^\otimes \neq I^\circ$, there are 2 possibilities. If $D_n \Subset I^\circ$, we have $|I^\otimes| = |I^\circ| - 1$ and $|K_{I^\otimes}| = |K_{I^\circ}| - 1$. If $|K_{I^\circ}| = |I^\circ|$, then $|K_{I^\otimes}| = |I^\otimes|$, but by the BC of the problem before the instantiation, 2. (prop. 1) holds. Otherwise, $|K_{I^\circ}| < |I^\circ|$, then $|K_{I^\otimes}| < |I^\otimes|$ and 1. (prop. 1) holds. If $D_n \subsetneq I^\circ$, we have $|I^\otimes| = |I^\circ| - 1$ and $|K_{I^\otimes}| = |K_{I^\circ}|$, and by BC, $|K_{I^\circ}| < |I^\circ|$. If $|K_{I^\circ}| < |I^\circ| - 1$, then $|K_{I^\otimes}| < |I^\otimes|$ and 1. (prop. 1) holds. Otherwise, $|K_{I^\circ}| = |I^\circ| - 1$, then $|K_{I^\otimes}| = |I^\otimes|$, and an inconsistency can only come from an intersection between $I^\otimes$ and at least one domain which is not included in $I^\otimes$. □

2.2 Some Probability Notions

Only the notions of discrete probability theory that are needed afterwards are presented here. For a good and complete introduction to probability theory, the reader can refer to [3].

Definition 2. *A* sample space Ω *is a non-empty set. An* event *on* Ω *is a subset of* Ω*. The set of all possible events is the set of parts of* Ω*, denoted* $\mathcal{P}(\Omega)$*. Two events are said* mutually exclusive *if and only if they have an empty intersection.*

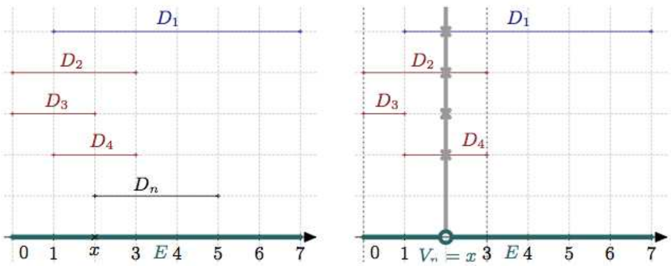


Fig. 1. When a variable (here V_n, on the left) is assigned to a value x (here $x = 2$, on the right), new Hall intervals may appear (here, the interval $\{0, 1, 3\}$ on the right). In this case, an `alldifferent` that was BC before the instantiation becomes inconsistent (here, domain D_1 should be reduced to $\{4..7\}$). Such *pre-Hall* intervals (here, on the left, interval $\{0, 1, 2, 3\}$) are described in Proposition 2.

For a given probabilistic phenomenon, the sample space is the set of all possible outcomes. For instance, consider the process of flipping a coin, the sample space would be the possible results of the coin flip, that is, the set $\{tails, heads\}$. In probability theory, events are what is observed and measured. In the example, an event could be $\{tails\}$, meaning that we want to observe the fact that the coin ended on tails. The events are defined as sets, but it is equivalent, and often more intuitive, to see them as predicates: "the result is tails". This event is for instance mutually exclusive with the event "the result is heads".

Definition 3. *A* probability distribution *is a function* $\mathbb{P}$ *from* $\mathcal{P}(\Omega)$ *to* $[0, 1]$*, such that* $\mathbb{P}[\Omega] = 1$*, and for any two mutually exclusive events* $A, B \subset \Omega$*,* $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$*. The triplet* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *is called a* probability space*.*

A probability distribution is a tool to measure events. In the previous example, we can define a probability distribution by: $\mathbb{P}[\{heads\}] = p$, $\mathbb{P}[\{tails\}] = 1 - p$ for a $p \in [0, 1]$. If $p = 0.5$, we have modeled the random process of flipping a perfectly fair coin. Other values of p model the same process with a biased coin.

Definition 4. *Two events* A *and* B *on a probability space* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *are said* independent *if and only if* $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$*. Otherwise, the events are said* correlated*.*

Intuitively, two events A and B are independent if A does not influence B, neither does B influence A.

Definition 5. *Let* A *and* B *be two events on a probability space* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$*. The* conditional probability of B, knowing A, *is* $\mathbb{P}[B|A] = \mathbb{P}[A \cap B]/\mathbb{P}[A]$*.*

The question here is to quantify the probability of B, under the condition that A *has* happened, that is, in the subspace of Ω where A happened. This subspace is simply A itself, measured by $\mathbb{P}[A]$. The probability of B within this subspace is $\mathbb{P}[A \cap B]$.

Definition 6. *Let* S *be a countable set and* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *a probability space. A* discrete random variable X *is a function* $X : \Omega \to S$*. The* distribution *of* X *is* $\mathbb{P}_X : S \to [0, 1]$ *such that* $\mathbb{P}_X[s] = \mathbb{P}[X^{-1}(s)]$ *for* $s \subset S$*.*

Proposition 3. *The distribution of a discrete random variable is a probability distribution.*

We do not detail the proof for Proposition 3 which can be found for instance in [3]. A discrete random variable is an event with a discrete or finite set of outcomes.

2.3 A probabilistic model for bound consistency

The constraint programmer's vocabulary includes expressions such as "this constraint `alldifferent` *has little chance* to filter here". The goal of this subsection is to propose a rigorous probabilistic model in which such expressions are well-defined. The key ingredients for consistency of an `alldifferent` constraint are the domains. Depending on their size and their relative position, the constraint may be consistent or not. However, sometimes it is not necessary to precisely know the domains in order to state if the constraint is consistent or not. For instance, an `alldifferent` constraint on three variables with domains of size 3 is always BC, whatever the exact positions of the domains. If the domains are of size 2, then the constraint may be BC or not, depending on the domains relative positions. But if E is also of size 2, the constraint is always inconsistent.

These basic remarks show that a partial knowledge on the domains and on E is sometimes sufficient to determine consistency properties. This is the basis of our probabilistic model: the domains become discrete random variables, with a given distribution (see in section 3). They are not fully determined: their exact sizes and positions are unknown. Only some macroscopic quantities are known.

Assume from now on that the domains $D_1...D_n$ are replaced by random variables with a known probability distribution. The sample space is naturally the set of subintervals of $E = \cup_i D_i$. In the following we make the assumption that the D_i variables are independent and that their distributions are identical. Let $\mathcal{D}$ be this distribution, with $\mathcal{D}(J) = \mathbb{P}[D_i = J]$ for $J \Subset E$ and $1 \leq i \leq n$. This hypothesis, as well as all the hypotheses that are made later, are summed up and discussed in section 4.1.

We are interested in studying how this probability evolves during the search process. At each node explored during the solving process, one could consider that a new problem is stated. This new problem would be the initial CSP where all the instantiated variables become constants, and the constraints are dealt with accordingly. From a probabilistic point of view, this is a loss of information, because the successive CSPs are very unlikely to be independent. In order to analyze precisely the probability of being BC, we propose to study the probability that an `alldifferent` constraint, already BC, *remains* BC after an instantiation. Such a reasoning could be called *conditional probability* but for simplicity reasons we just call about "probability". This probability can be computed from Proposition 2.

Proposition 4. *For an `alldifferent` constraint on variables $V_1...V_n$, where the domains are random variables with probability distribution $\mathcal{D}$ on an interval E of size m, the probability $P_{m,n,v}$ of remaining BC after instantiating the variable V_n to a value x at position v is $P_{m,n,v} =$*

$$\prod_{l=1}^{n-2} \prod_{\substack{\alpha = max(0, \\ v-l)}}^{\substack{min(v, \\ m-l-1)}} \left(1 - (1 - r_{m,l+1,\alpha,v}) \binom{n-1}{l} p_{m,l+1}^{l} \left((1-p_{m,l+1})^{n-l-1} - q_{m,l+1,\alpha}^{n-l-1}\right)\right)$$

where, for an interval $I \Subset E$, of size l, at position $\alpha = min(I) - min(E)$,

$p_{m,l} = \mathbb{P}[D_i \subset I]$,

$q_{m,l,\alpha} = \mathbb{P}[D_i \cap I = \emptyset]$,
$r_{m,v,l,\alpha} = \mathbb{P}[D_i \subset I | x \in D_i]$.

and these $I \Subset E$ are assumed to be independent.

Proof. The constraint remains BC iff $\forall I^{\otimes} \Subset E^{\otimes}$ such that $I^{\otimes} \neq I^{\circ}$, at least one of the conditions (1), (2) of Proposition 2 is false. Because the $I^{\otimes}$ are assumed to be independent, "remaining BC" happens with probability:

$$\prod_{\substack{I^{\otimes} \Subset E^{\otimes} \\ I^{\otimes} \neq I^{\circ}}} (1 - \mathbb{P}[(1)]\mathbb{P}[(2)])$$

The subintervals of E are represented by their size l and position α. For a given size l, the condition $I^{\otimes} \neq I^{\circ}$ restricts the range for α to $\{max(0, v-l)...min(v, m-l-1)\}$. By definition, $\mathbb{P}[(1)] = 1 - r_{m,l+1,\alpha,v}$. Consider now the probability for (2). For l given domains, the probability that they are included in I° is $p^l_{m,l+1}$, and there are $\binom{n-1}{l}$ possibilities for choosing these domains among the $n-1$ domains. Consider now the $n-l-1$ other domains: they must not be totally included in $I^{\otimes}$, which happens with probability $(1-p_{m,l+1})^{n-l-1}$. In addition, one must forbid the case where they would all be totally outside $I^{\otimes}$, which happens with probability $q^{n-l-1}_{m,l+1,\alpha}$. Hence the formula. $\square$

The "atomic probabilities" $p_{m,l}$, $q_{m,l,\alpha}$ and $r_{m,v,l,\alpha}$ depend on the distribution $\mathcal{D}$ chosen for the domains (and their parameters of course). If they can be computed for a given distribution, then the formula gives the desired probability. Note that the formula relies on two products, each with n terms. Without further knowledge on the expressions of $p_{m,l}$, $q_{m,l,\alpha}$ and $r_{m,v,l,\alpha}$, the complexity of computing $P_{m,n,v}$ has an order of magnitude of $O(n^2)$.

3 Computing the probabilities

In this section, we study two cases for the domains distribution $\mathcal{D}$. For each of them, the atomic probabilities can be computed and inserted in Proposition 4. We still have to deal with the complexity of the formula, which has an order of magnitude at least n^2. It is thus interesting to consider the asymptotic expansion of the probability when m tends towards infinity. The asymptotic expansion gives useful (and computationally cheap) numerical values for the probability of remaining BC.

3.1 Uniform domain distribution

In order to model the general case, that is, when the domains distribution is not biased toward specific domains (of a given lower bound, upper bound, etc), we define a uniform distribution for D_i.

Definition 7. *Given an interval E of size m, a random variable D_i on $\{J, J \Subset E\}$, D_i has a* uniform distribution *if and only if the probability $\mathbb{P}[D_i = J]$ does not depend on J.*

Remark 1. This implies:

- $\forall J \Subset E, \mathbb{P}[D_i = J] = \frac{1}{|\{J', J' \Subset E\}|} = \frac{2}{m(m-1)}$,
- $\mathbb{P}[|D_i| = d] = \frac{2(m-d+1)}{m(m-1)}$, thus the distribution of the domains sizes is biased toward small sizes.

With this distribution, the atomic probabilities $p_{m,l}$, $q_{m,l,\alpha}$ and $r_{m,v,l,\alpha}$ of the formula in Proposition 4 can be computed.

Lemma 1. *With the notations of Proposition 4, assume that $\mathcal{D}$ is a uniform distribution on E. Then for an interval $I \Subset E$, of size l, at position α:*

$p_{m,l} = \frac{l(l-1)}{m(m-1)}$;
$q_{m,l,\alpha} = \frac{\alpha(\alpha-1)+(m-l-\alpha)(m-l-\alpha-1)}{m(m-1)}$;
$r_{m,v,l,\alpha} = \frac{(l-1)+(l-1+\alpha-v)(v-\alpha)}{(m-1)+(m-1-v)v}$.

Proof. From Proposition 4, we have $p_{m,l} = \mathbb{P}[D_i \Subset I]$. Considering the domains D_i according to their size leads to:

$$\begin{aligned}
\mathbb{P}[D_i \Subset I] &= \sum_{d=2}^{m} \mathbb{P}[(D_i \Subset I) \wedge (|D_i| = d)] \\
&= \sum_{d=2}^{m} \mathbb{P}[(D_i \Subset I)|(|D_i| = d)]\mathbb{P}[|D_i| = d] \\
&= \sum_{d=2}^{m} \mathbb{P}[(D_i \Subset I)|(|D_i| = d)]\tfrac{2(m-d+1)}{m(m-1)} \quad \text{from Remark 1} \\
&= \sum_{d=2}^{m} \tfrac{l-d+1}{m-d+1}\tfrac{2(m-d+1)}{m(m-1)} \\
&= \tfrac{l(l-1)}{m(m-1)}.
\end{aligned}$$

For $q_{m,l,\alpha} = \mathbb{P}[D_i \cap I = \emptyset]$, the property is equivalent to: $D_i \subset E \backslash I$. The formula is thus similar to the computation for $p_{m,l}$, by considering two subintervals of E, one strictly on the left of I and one strictly on the right.

Finally, we have $r_{m,v,l,\alpha} = \mathbb{P}[D_i \Subset I | x \in D_i] = \frac{\mathbb{P}[D_i \Subset I \wedge x \in D_i]}{\mathbb{P}[x \in D_i]}$. Thus, $r_{m,v,l,\alpha} = \frac{\sum_{d=2}^{m} \mathbb{P}[(D_i \Subset I \wedge x \in D_i)|(|D_i|=d)]\mathbb{P}[|D_i|=d]}{\sum_{d=2}^{m} \mathbb{P}[x \in D_i|(|D_i|=d)]\mathbb{P}[|D_i|=d]}$. The events $\{x \in D_i | \{|D_i| = d\}\}$ and $\{\{D_i \Subset I \wedge x \in D_i\} | \{|D_i| = d\}\}$ are respectively equivalent to $\{D_i \Subset E \cap [x-d+1, x+d-1]\}$ and $\{D_i \Subset I \cap [x-d+1, x+d-1]\}$. From there, the computation is similar to the one of $p_{m,l}$. □

Assuming we know the quantities $p_{m,l}$, $q_{m,l,\alpha}$ and $r_{m,v,l,\alpha}$, Proposition 4 gives the probability of remaining BC. The following theorem gives an asymptotic expansion of this formula, when m tends toward infinity while the ratios n/m and $v/(m-1)$ remain constant. Depending on the values of these ratios, different behaviors appear.

Theorem 1. *Consider an `alldifferent` constraint on variables $V_1 ... V_n$, with domains $D_1 ... D_n$, such that the constraint is BC. Assume that the variable V_n is set to a value $x \in D_n$, and let v be the position of x within E ($v = x - min(E)$). Let $\rho = n/m$ and $\nu = v/(m-1)$. Assume that the domains D_i are uniformly distributed on E. Let $P_{m,n,v}$ be the probability that the `alldifferent` constraint remains BC after the instantiation. Then, when $m \to \infty$ with fixed ρ and ν:*

$$P_{m,\rho m,\nu(m-1)} = \begin{cases} 1 - \frac{4\rho}{m} + O\left(\frac{1}{m^2}\right) & \text{if } 0 < \rho < 1 \text{ and } 0 < \nu < 1 \\ 1 - \frac{2\rho(1-e^{-4\rho})}{m} + O\left(\frac{1}{m^2}\right) & \text{if } 0 < \rho < 1 \text{ and } \nu = 0 \text{ or } 1 \\ 1 - \frac{2(1-e^{-4})+S}{m} + O\left(\frac{1}{m^2}\right) & \text{if } \rho = 1 \text{ and } \nu = 0 \text{ or } 1 \\ 1 - \frac{4+\frac{S'}{2\nu(1-\nu)}}{m} + O\left(\frac{1}{m^2}\right) & \text{if } \rho = 1 \text{ and } 0 < \nu < 1 \end{cases}$$

where S and S' are two constants that can be numerically computed:

$$S = \sum_{k=1}^{\infty} \frac{2^k e^{-2k} k^k}{(k-1)!} \approx 1.94264 \quad and \quad S' = \sum_{k=1}^{\infty} \frac{2^k e^{-2k} k^k (k+1)}{(k-1)!} \approx 11.9359.$$

The proof is detailed in Appendix A. It comes from using Lemma 1 in Proposition 4, and carefully bounding each term of the sum. Note that all these asymptotic expansions depend on the values of ρ and ν. However, they are always very close to 1. The first order term is known, with a second-order error term.

Corollary 1. *The probability that an `alldifferent` constraint, in the uniform model, remains BC after an instantiation*

- *can be approximated, at the order $O(\frac{1}{m^2})$, in time $O(1)$,*
- *depends on $\rho = n/m$, and $\nu = v/(m-1)$, that is, the number of available values for all the variables and the position of the instantiated value.*

This corollary can be used in two ways. Firstly, it gives an approximation, in the limit $m \to \infty$, with fixed ρ and ν, of the probability of remaining BC. This approximation can be computed in constant time, thus being competitive with any consistency algorithm. The precision of this approximation, of order $\frac{1}{m^2}$, is experimentally studied in Section 4.3. Secondly, it exhibits several different behaviors depending on the values of ρ and ν. Obviously, these quantities are meaningful in determining the behavior of an `alldifferent` constraint. This will be discussed in Section4.2.

3.2 Fixed-size domain distribution

It can also happen that the domains of a CSP are biased. Most of the time, they are biased toward a given size: for instance, many puzzle problems are stated on n variables with domains $\{1...n\}$ (magic squares, sudoku, latin squares, etc). In order to model this case, we define a so-called fixed-size uniform distribution:

Definition 8. *Given an interval E of size m, a fixed-size distribution of parameter d is an uniform distribution on $\{J' \Subset E, |J'| = d\}$.*

Remark 2. This implies:

- $\mathbb{P}[D_i = J] = \frac{\delta_{|J|,d}}{m-d+1}$, δ is the Kronecker delta: for a, b integers, $\delta_{a,b} = 1$ if $a = b$, 0 otherwise.
- $\mathbb{P}[|D_i| = d'] = \delta_{d,d'}$, thus only domains of size d are kept.

For this distribution, we add, in the notations, d as a parameter of the atomic probabilities: $p_{m,l}(d)$, $q_{m,l,\alpha}(d)$ and $r_{m,v,l,\alpha}(d)$. The atomic probabilities are given by the following lemma :

Lemma 2. *With the notations of Proposition 4, assume that $\mathcal{D}$ is a fixed-size distribution of parameter d on E. Then for an interval $I \Subset E$, of size l, at position α:*

$p_{m,l} = max\left(0, \frac{l-d+1}{m-d+1}\right)$;
$q_{m,l,\alpha} = max\left(0, \frac{\alpha-d+1}{m-d+1}\right) + max\left(0, \frac{m-l-\alpha-d+1}{m-d+1}\right)$;
$r_{m,v,l,\alpha} = \mathcal{H}(l-d)\frac{min(v,\alpha+l-d)-max(\alpha,v-d+1)+1}{min(v,m-d)-max(0,v-d+1)+1}$. *$\mathcal{H}$ is the Heaviside function: $\mathcal{H}(x) = 1$ if $x \geq 0$ and 0 otherwise.*

Proof. The proof is exactly the same as for the uniform distribution, replacing $\mathbb{P}[|D_i| = k]$ by $\delta_{k,d}$. $\square$

Assuming we know the atomic probabilities, we can compute the probability of remaining BC from Proposition 4. The following theorem gives an asymptotic expansion of this formula, when m tends toward infinity while the ratios $\rho = n/m$, $\nu = v/(m-1)$ and $\mu = d/m$ remain constant.

Theorem 2. *Consider an* `alldifferent` *constraint on variables $V_1...V_n$, with domains $D_1...D_n$, such that the constraint is BC. Assume that the variable V_n is set to a value $x \in D_n$, and let v be the position of x within E ($v = x - min(E)$). Let $\rho = n/m$ and $\nu = v/(m-1)$. Assume that the domains D_i have a fixed-size distribution of parameter d on E, and let $\mu = d/m$. Let $P_{m,n,v}(d)$ be the probability that the* `alldifferent` *constraint remains BC after the instantiation. Then, when $m \to \infty$, and for fixed ρ, ν and μ:*

$$P_{m,\rho m,\nu(m-1)}(\mu m) = \begin{cases} 1 & \text{if } \nu = 0 \text{ or } \nu = 1 \text{ or } \mu \geq \rho \\ 1 - o\left(\frac{1}{m^2}\right) & \text{if } \rho < 1 \text{ and } 0 < \nu < 1 \\ 1 - \frac{T_{\nu\mu}}{m\Phi_{\nu\mu}} + O\left(\frac{1}{m^2}\right) & \text{if } \rho = 1 \text{ and } 0 < \nu < 1 \end{cases}$$

where $\Phi_{\nu\mu} = min(\nu, 1-\mu) - max(0, \nu - \mu)$ and

$$T_{\nu\mu} = \begin{cases} f(\mu) & \text{if } \nu < \mu \text{ and } \nu > 1-\mu \\ \frac{f(\mu)}{2} + \delta_{\nu,\mu}\left(\frac{f(\mu)}{2} - (1-\mu)g(\mu)\right) & \text{if } \nu \geq \mu \text{ and } \nu > 1-\mu \\ \left(\frac{f(\mu)}{2} - (1-\mu)g(\mu)\right)(\delta_{\nu,\mu} + \delta_{\nu,1-\mu}) & \text{if } \nu \geq \mu \text{ and } \nu \leq 1-\mu \\ \frac{f(\mu)}{2} + \delta_{\nu,1-\mu}\left(\frac{f(\mu)}{2} - (1-\mu)g(\mu)\right) & \text{if } \nu < \mu \text{ and } \nu \leq 1-\mu \end{cases}$$

with $g(\mu) = \sum_{k=1}^{\infty} \left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{(k-1)!}$ *and* $f(\mu) = \sum_{k=1}^{\infty} (k+1)\left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{(k-1)!}$.

Again, the proof is detailed in Appendix B. It is very similar to the that of Theorem1. In this model, the asymptotic expansion depends on ρ and ν and also on the parameter μ. We can remark that the asymptotic expansions are closer to 1 than in the uniform model in some cases.

Corollary 2. *The probability that an* `alldifferent` *constraint, in the fixed-size model, remains BC after an instantiation*

- *can be approximated, at the order $O(\frac{1}{m^2})$, in time $O(1)$,*
- *depends on $\rho = n/m$, $\nu = v/(m-1)$ and $\mu = d/m$, that is, the number of available values for all the variables, the position of the instantiated value, and the common size of the domains.*

The same remarks as for Corollary 1 apply, and again this result will be discussed in Section4.

4 Analysis and Experiments

This section discusses the contributions of this paper, from several point of views: validity of the model, analysis of the two theorems, and experimental evaluation.

4.1 Validity of the Hypothesis

We propose a probabilistic *model* for BC of an `alldifferent` constraint. Like all models, it relies on a simplistic view of the reality, and this simplification is expressed by the mathematical hypotheses that have been used.

The first hypothesis is on the domains, considered as random variables that are independent and identically distributed. As long as we do not consider other constraints, the independence assumption is reasonable, because the variables of the constraint do not interact outside the constraint under study. Going one step further, one could also consider the interactions between the constraints, for instance several `alldifferent` constraints as studied in [1]. This is part of our further work. For the identical distribution, this is clearly a loss of information on the domains. Actually, loosing information on the domains is precisely the key point of the probabilistic model.

Another hypothesis holds on the subintervals of E used in Proposition 4. The $I \Subset E$ are assumed to be independent; more precisely, conditions (1) and (2) of Proposition 2 are assumed to be independent when I varies. These conditions express the "Hall status" of these intervals. The main argument to justify their independence is that this Hall property is local, that is, it cannot be transferred from an interval I to another. Given an interval $I \Subset E$ that is Hall for instance, there is no particular reason why other intervals $I' \Subset E$ should be Hall. The independence hypothesis seems thus reasonable.

The important results within the probabilistic model are the two theorems of Section 3, and they are valid in the limit $m \to \infty$. This asymptotical result, which is rigorous, yields an approximated computation for $P_{m,n,v}$, known until the second order error term. In practice, these theorems are valid for large values of m (at the beginning of the search process).

The last point is not a hypothesis, but also a limitation of the model: only two distributions for the domains have been considered. These distributions correspond to extreme scenarios: the uniform distribution implies that all the possible subintervals of E are equally weighted, that is, there is no particular bias in the domains distribution. On the contrary, the fixed-size uniform distribution restricts the domains to be of a fixed size, as it happens quite often, in problems coming from Artificial Intelligence for instance. To our knowledge, this should be sufficient to model quite a lot of real-life scenarios.

4.2 Technical Analysis

We propose the following terminology to discuss the Theorems 1 and 2. The number of variables over the total number of available values for them, $\rho = n/m$, determines if the problem is sharp ($\rho = 1$, exactly as many variables as values) or not ($\rho < 1$, some values will not be used in the end). The position of the instantiation $\nu = v/m$ is another important parameter. For $\nu = 0$ or $\nu = 1$, the instantiation is performed at a bound of E, which is possible, with a variable-value heuristics that chooses first a variable with a domain at a bound of E, say the minimum, then the minimum value of this domain. Note that there is a symmetry when ν is replaced by $1 - \nu$. Finally, if the domains size are fixed at d, then $\mu = d/m$ is the ratio of the domains size over the number of available values. A small value for μ mean that the domains are tight. If $\mu = 1$, then the problem is not tight, and all the domains equal E.

Figure 2 shows the plots of the several limits given by Theorem 1. Problems that are not sharp ($\rho < 1$) have few chances to actually filter. As soon as $m \geq 50$, the probability is higher than 0.9, meaning that there are at most 10% chances to remove a value (or more) from the domains. Notice that the curves are decreasing in ρ: sharper problems have better chances to filter. Instantiating

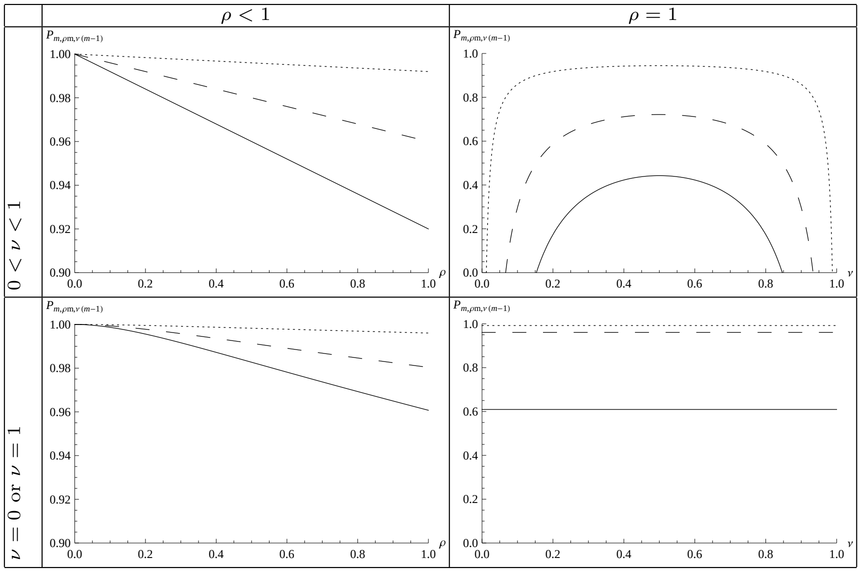


Fig. 2. The asymptotic expansions of theorem 1, depending on the values of ρ and ν. The plain curves are for $m = 50$, the large dashed curves for $m = 100$ and the small dashed curves for $m = 500$. The plot range is $[0, 1]$ for the first line, and is restricted to $[0.9, 1]$ for the second line (otherwise the curves are too close to be identified).

at a bound (bottom left) yields a smaller probability to actually filter. For sharp problems, the behavior strongly depends on the value of ν.

The fixed-size model of Theorem 2 is quite different, as seen on Figure 3. For problems that are not sharp, the asymptotic expansion is 1 (not plotted): the domains being very specific, Hall intervals appear less frequently. For instance, the smallest possible Hall interval is of size d (2 in the uniform model). For sharp problems, the asymptotic expansion is of order $1/m$, and depends on the tightness. First, note that the sizes of the problem, m, used to the left and right figures are different. Indeed, both functions $f(\mu)$ and $g(\mu)$ do not converge when $\mu \to 0$ and the validity of the approximation is not reached at the same union size m. On both figures, there are clearly 3 different phases, as stated by Theorem 2 for sharp problems. The probability goes through the 3 conditions, depending on the relative positions of ν, μ and $1 - \mu$. In addition, the behaviors are different depending on the domains size. In general, in order to filter more, it is preferable to instantiate at a bound.

4.3 Practical Evaluation

We have tested the validity of the asymptotic expansions of Theorems 1 and 2, with the BC algorithm for `alldifferent` provided by Choco [6]. For the experiments on Theorem 1 (resp. 2), given the parameters n and m (resp. n, m and d), we randomly generate n domains on the interval $\{1...m\}$ according to the uniform distribution (resp. fixed-size uniform distribution of parameter d). If the generated domains do not cover E (which may happen for small values of n), then the

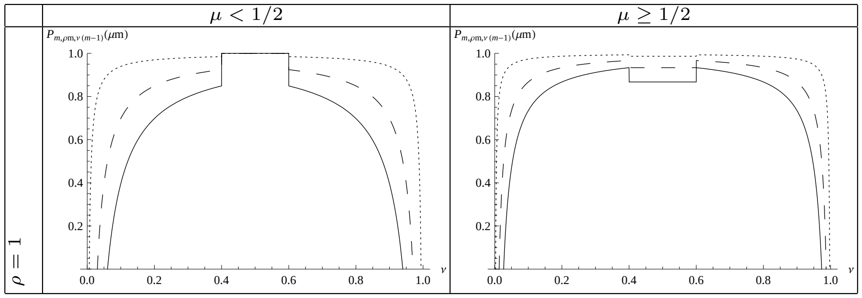


Fig. 3. The asymptotic expansions of theorem 2, for $\rho = 1$. On the left, $\mu = 0.4$, and on the right, $\mu = 0.6$. The plain curves are for $m = 50$ (resp 500), the large dashed curves for $m = 100$ (resp 1000) and the small dashed curves for $m = 500$ (resp 5000).

experiment for these domains is removed.

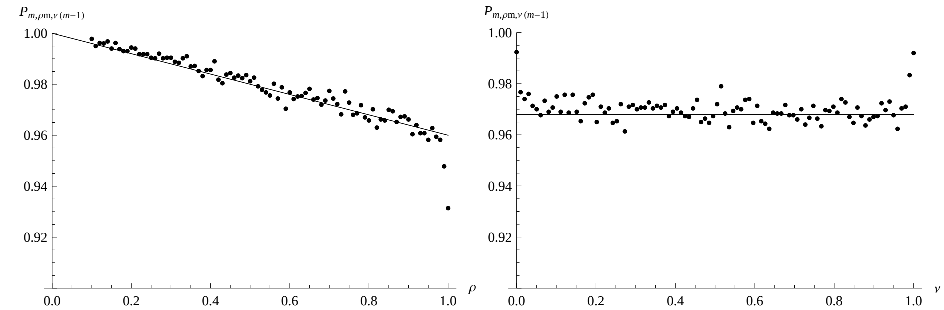
An `alldifferent` constraint is then declared on n variables, with the randomly generated domains. The BC algorithm is called a first time, then an instantiation at position v is performed, v being a parameter of the test. Forward checking is applied on the binary difference constraints. This removes the obvious inconsistent values. We apply again the BC algorithm and check if the new filtering has removed at least one value (negative test), or no value at all (positive). Each test is repeated 3000 times for each parameter values. The proportion of positive and negative tests on these 3000 tests is represented as a dot on the plots (Figure 4). The theoretical curves are drawn with plain line.

We see on the plots that the dotted graphics are very close to the theoretical curves (Figures 4(a) and 4(b)). Very quickly (for $m > 20$), the error is around 1% (Figure 4(c)). These results experimentally validate the hypotheses of Section 4.1, and confirm the theoretical ones.

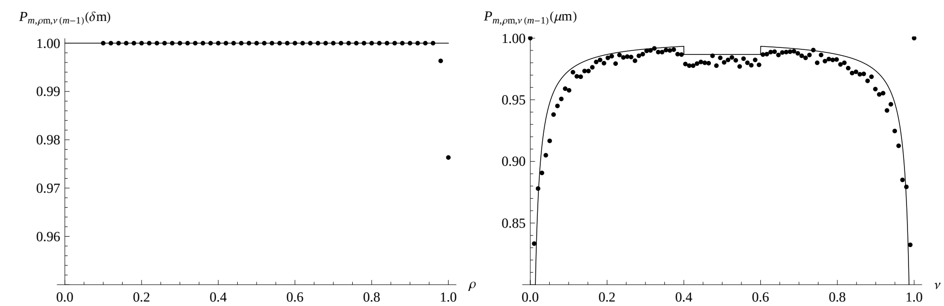
5 Conclusion

We have presented in this paper an accurate probabilistic indicator for the bound consistency of an `alldifferent` constraint. Moreover, this indicator can be computed in constant time. Not surprisingly, we still have to pay price; the intricate computations required to derived the easy-to-use expressions of Theorems 1 and 2. In a few words, this paper shows that, considering an `alldifferent` constraint alone, bound consistency has a limited efficiency. On real-life problems, this leads us to think that its efficiency lies in its interactions with the other constraints.

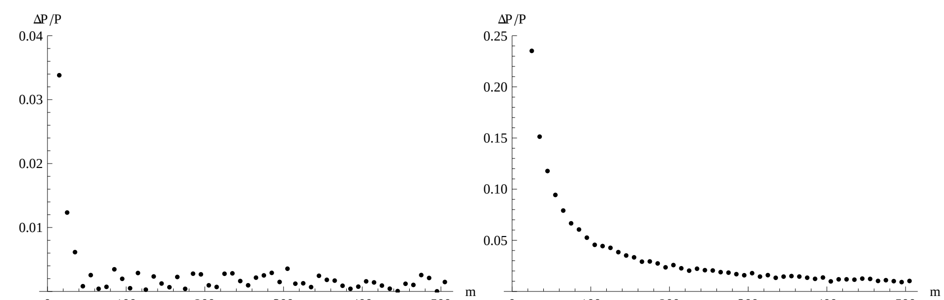
However, our indicator allows us to tackle some challenges related to the practical uses of global constraints, *e.g.*, triggering the filtering algorithms and guiding the heuristics. The filtering algorithms for global constraints are not effective all the time, and they may have a high complexity in some cases. We can prevent this phenomenon by limiting the practical complexity of the filtering by using the probabilistic indicator, one could inhibit most of the useless filtering. At another level, the consistency behavior differs, depending on some heuristic choices (variable value selection). The probabilistic model can thus help in designing heuristics on high level criteria (*e.g.*, the ν parameter).



(a) Uniform model with $m = 100$. On the left, ρ varies, and $\nu = 0.4$. On the right, ν varies, and $\rho = 0.8$.



(b) Fixed size uniform model. On the left, $m = 100$, ρ varies, $\nu = 0.4$, and $\mu = 0.5$. On the right, $m = 500$, ν varies, $\rho = 1$ and $\mu = 0.6$.



(c) $\Delta P/P$ represents the relative difference between theoretical and measured probabilities. On the left, uniform model with $\rho = 0.7$ and $\nu = 0.3$. On the right, fixed-size model with $\rho = 1$, $\nu = 0.4$ and $\mu = 0.6$.

Fig. 4. Theoretical (plain curves) and measured (dotted plots) probabilities

Finally, which global constraints are good candidates for such a probabilistic approach? The global cardinality constraint and the `NValue` constraint have a combinatorial structure close to that of `alldifferent`. We hope to present a similar study of these constraints in further works.

References

1. C. Bessiere, G. Katsirelos, N. Narodytska, C.-G. Quimper, and T. Walsh. Propagating conjunctions of alldifferent constraints. In *AAAI*, 2010.
2. C. Bessière and P. van Hentenryck. To Be or Not to Be... a Global Constraint. In *Principles and Practice of Constraint Programming*, pages 789–794, 2003.
3. G.R. Grimmett and D.R. Stirzaker. *Probability and Random Processes (second edition)*. Oxford Science Publications, 1992.
4. I. Katriel. Expected-case analysis for delayed filtering. In *CPAIOR*, volume 3990 of *Lecture Notes in Computer Science*, pages 119–125. Springer, 2006.
5. J.-F. Puget. A fast algorithm for the bound consistency of alldiff constraints. In *AAAI/IAAI*, pages 359–366, 1998.
6. Choco Team. Choco: an open source java constraint programming library. Research report 10-02-INFO, Ecole des Mines de Nantes, 2010.
7. W.J. van Hoeve. The alldifferent constraint: A survey. *CoRR*, cs.PL/0105015, 2001.
8. A. Zanarini and G. Pesant. Solution counting algorithms for constraint-centered search heuristics. *Constraints*, 14(3):392–413, 2009.

A Proof of theorem 1

The probability $P_{m,n,v}$, that an `alldifferent` constraint on n variables(with union of size m) stays consistent after the instanciation of a variable at the value x (at position v) is given by proposition 4. We set $L_{m,n,v,l,\alpha} = (1-r_{m,v,l+1,\alpha})\binom{n-1}{l}p^l_{m,l+1}\left((1-p_{m,l+1})^{n-l-1} - q^{n-l-1}_{m,l,\alpha}\right)$. Thus,

$$\begin{aligned}
P_{m,n,v} &= \prod_{l=1}^{n-2}\prod_{\alpha=max(0,v-l)}^{min(v,m-l-1)} (1-L_{m,n,v,l,\alpha}) \\
&= \exp\left(\sum_{l=1}^{n-2}\sum_{\alpha=max(0,v-l)}^{min(v,m-l-1)} \log(1-L_{m,n,v,l,\alpha})\right)
\end{aligned}$$

We also define $X_{m,n,v,l} = \sum_{\alpha=max(0,v-l)}^{min(v,m-l-1)} \log(1-L_{m,n,v,l,\alpha})$, such that $P_{m,n,v} = \exp\left(\sum_{l=1}^{n-2} X_{m,n,v,l}\right)$.

Let's set $\rho = n/m$ and $\nu = v/(m-1)$. We have $0 < \rho \leq 1$ and $0 \leq \nu \leq 1$. We distinguish the 4 following cases : $\rho < 1$ and $0 < \nu < 1$; $\rho < 1$ and $\nu = 0$ or 1; $\rho = 1$ and $0 < \nu < 1$; $\rho = 1$ and $\nu = 0$ ou 1. In each case, we show that most of the terms are negligibles.

A.1 Case $\rho < 1$ and $0 < \nu < 1$

Let $0 < \epsilon < 1/2$, we cut the sum on l as follows : $l = 1, 2, 3$; $3 < l < (n-1)^\epsilon$; $(n-1)^\epsilon \leq l \leq n-2$.

- $l = 1, 2, 3$. We have

$$\begin{aligned}
X_{m,n,v,1} &= \sum_{\alpha=max(0,v-1)}^{min(v,m-2)} \log(1-L_{m,n,v,1,\alpha}) \\
&= \sum_{\alpha=v-1}^{v} \log(1-L_{m,n,v,1,\alpha}) \\
&= \log(1-L_{m,n,v,1,v-1}) + \log(1-L_{m,n,v,1,v}).
\end{aligned}$$

The asymptotics are given by $L_{m,\rho m,\nu(m-1),1,\nu(m-1)-1} = \frac{2\rho}{m} + O\left(\frac{1}{m^2}\right)$ and $L_{m,\rho m,\nu(m-1),1,\nu(m-1)} = \frac{2\rho}{m} + O\left(\frac{1}{m^2}\right)$. Thus

$$X_{m,\rho m,\nu(m-1),1} = -\frac{4\rho}{m} + O\left(\frac{1}{m^2}\right). \tag{1}$$

In the same way, we can see that

$$X_{m,\rho m,\nu(m-1),2} = O\left(\frac{1}{m^2}\right) \text{ and } X_{m,\rho m,\nu(m-1),3} = O\left(\frac{1}{m^3}\right). \tag{2}$$

– $3 < l < (n-1)^\epsilon$. As $r_{m,v,l+1,\alpha} \leq 1$ and $q_{m,l+1,\alpha} \leq 1 - p_{m,l+1} \leq 1$, we have $L_{m,n,v,l,\alpha} \leq U_{m,n,l} := \binom{n-1}{l}\left(\frac{l(l+1)}{m(m-1)}\right)^l$.

$$\begin{aligned}
\frac{U_{m,n,l+1}}{U_{m,n,l}} &= \frac{(n-l-1)(l+2)}{m(m-1)}\left(\frac{l+2}{l}\right)^l \\
&\leq \frac{(n-1-(n-1)^\epsilon)((n-1)^\epsilon+2)}{m(m-1)}e^2 \qquad \text{because } 0<\epsilon<1/2 \\
&\leq \frac{n-1}{m-1}\frac{(n-1)^\epsilon}{m}\left(1+\frac{2}{(n-1)^\epsilon}\right)e^2 \\
&\leq \frac{\rho^{1+\epsilon}}{m^{1-\epsilon}}\left(1+2^{1-\epsilon}\right)e^2 \qquad \text{because } n \geq 3 \\
&\leq 3e^2\frac{\rho^{1+\epsilon}}{m^{1-\epsilon}}
\end{aligned}$$

It shows that at least $\forall m \geq m_0 = (3e^2\rho^{1+\epsilon})^{1/(1-\epsilon)}$, U decreases in the "l" direction. Thus for $l \geq 4$, $L_{m,n,v,l,\alpha} \leq U_{m,n,4} = \binom{n-1}{4}\frac{4^4 5^4}{m^4(m-1)^4}$. As $U_{m,n,4} \to 0$ when $m \to \infty$, $\exists m_1$ such that $\forall m \geq m_1$, $L_{m,n,v,l,\alpha} \leq 1/2$, and so $\log(1 - L_{m,n,v,l,\alpha}) \geq -2L_{m,n,v,l,\alpha}$. Then, $\forall m \geq m_1$,

$$\begin{aligned}
\left|\sum_{l=4}^{(n-1)^\epsilon-1} X_{m,n,v,l}\right| &\leq 2\sum_{l=4}^{(n-1)^\epsilon-1}\ \sum_{\alpha=max(0,v-l)}^{min(v,m-l-1)} U_{m,n,4} \\
&\leq \frac{4^4 5^4}{12}\frac{\rho^{4+\epsilon}}{m^{3-\epsilon}}
\end{aligned}$$

which means that

$$\sum_{l=4}^{(n-1)^\epsilon-1} X_{m,n,l} = O\left(\frac{1}{m^{3-\epsilon}}\right) = o\left(\frac{1}{m^2}\right). \tag{3}$$

– $(n-1)^\epsilon \leq l \leq n-2$. As $r_{m,v,l+1,\alpha} \leq 1$ and $q_{m,l+1,\alpha} \leq 1-p_{m,l+1}$, we have $L_{m,n,v,l,\alpha} \leq V_{m,n,l} := \binom{n-1}{l}p_{m,l+1}^l(1-p_{m,l+1})^{n-l-1}$.

$$\begin{aligned}
\frac{V_{m,n+1,l}}{V_{m,n,l}} &= \frac{n}{n-l}\left(1-\frac{l(l+1)}{m(m-1)}\right) \\
&\geq \frac{n}{n-l}\left(1-\frac{l(l+1)}{n(n-1)}\right) \\
&= 1+\frac{l(n^2-n(l+2))}{n(n-1)(n-l)} \\
&\geq 1 \qquad \text{because } l \leq n-2
\end{aligned}$$

It shows that V increases in the "n" direction, thus $V_{m,n,l} \leq V_{m,m,l}$. Using $\sqrt{2\pi k} k^k e^{-k} \leq k! \leq \sqrt{2\pi k} k^k e^{-k} e^{\frac{1}{12k}}$ and $m \geq 3$, we have :

$$\begin{aligned}
\binom{m-1}{l} &\leq \frac{1}{\sqrt{2\pi}} \sqrt{\frac{m-1}{l(m-l-1)}} \frac{(m-1)^{m-1}}{l^l (m-l-1)^{m-l-1}} e^{\frac{1}{24}} \\
&= \frac{e^{\frac{1}{24}}}{\sqrt{2\pi}} \frac{1}{\sqrt{m-1}} \frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}} \frac{1}{\left(\frac{l}{m-1}\right)^l \left(1-\frac{l}{m-1}\right)^{m-l-1}}
\end{aligned} \tag{4}$$

Besides, we have

$$\begin{aligned}
\left(\frac{l(l+1)}{m(m-1)}\right)^l &= \left(\frac{l}{(m-1)}\right)^{2l} \left(1+\frac{1}{l}\right)^l \left(1-\frac{1}{m}\right)^l \\
&\leq e \left(\frac{l}{(m-1)}\right)^{2l}
\end{aligned}$$

and

$$\begin{aligned}
\left(1-\frac{l(l+1)}{m(m-1)}\right)^{m-l-1} &= \left(1-\left(\frac{l}{m-1}\right)^2\right)^{m-l-1} \left(1+\frac{1}{m+l-1}\right)^{m-l-1} \left(1-\frac{1}{m}\right)^{m-l-1} \\
&\leq \left(1-\left(\frac{l}{m-1}\right)^2\right)^{m-l-1} e^{(m-l-1)\log\left(1+\frac{1}{m+l-1}\right)} \\
&\leq \left(1-\left(\frac{l}{m-1}\right)^2\right)^{m-l-1} e^{\frac{m-l-1}{m+l-1}} \\
&\leq e \left(1-\left(\frac{l}{m-1}\right)^2\right)^{m-l-1}.
\end{aligned}$$

Then

$$V_{m,m,l} \leq \frac{e^2 e^{\frac{1}{24}}}{\sqrt{2\pi}} \frac{1}{\sqrt{m-1}} \frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}} \left(\left(\frac{l}{m-1}\right)^{\frac{l}{m-1}} \left(1+\frac{l}{m-1}\right)^{1-\frac{l}{m-1}}\right)^{m-1}$$

We can see that $\left(\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)\right)^{-\frac{1}{2}}$ is maximal when $l = (n-1)^\epsilon$ (at least for n sufficiently large), that is

$$\begin{aligned}
\frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}} &\leq \frac{1}{\sqrt{\frac{(n-1)^\epsilon}{m-1}\left(1-\frac{(n-1)^\epsilon}{m-1}\right)}} \\
&\leq \frac{1}{\sqrt{\frac{(n-1)^\epsilon}{n-1}\frac{n-1}{n}\frac{n}{m}\frac{m}{m-1}}} \frac{1}{\sqrt{1-\frac{1}{(m-1)^{1-\epsilon}}}} \\
&\leq (n-1)^{\frac{1}{2}-\frac{\epsilon}{2}} \frac{1}{\sqrt{\frac{2}{3}\rho}} \frac{1}{\sqrt{1-\frac{1}{\sqrt{2}}}} \\
&\leq \sqrt{6}\rho^{-\frac{\epsilon}{2}} (m-1)^{\frac{1}{2}-\frac{\epsilon}{2}}
\end{aligned} \tag{5}$$

Moreover, $\exists x_0 \in]0,1[$ such that the function $x \mapsto x^x(1+x)^{1-x}$, decreases on $[0,x_0]$ and increases on $]x_0,1]$. Thus, $\left(\frac{l}{m-1}\right)^{\frac{l}{m-1}}\left(1+\frac{l}{m-1}\right)^{1-\frac{l}{m-1}}$ is maximal when $l=(n-1)^\epsilon$ or when $l=(n-2)$. We have :

$$\begin{aligned}
\left(\left(\frac{(n-1)^\epsilon}{m-1}\right)^{\frac{(n-1)^\epsilon}{m-1}}\left(1+\frac{(n-1)^\epsilon}{m-1}\right)^{1-\frac{(n-1)^\epsilon}{m-1}}\right)^{m-1} &= \left(\frac{(n-1)^\epsilon}{m-1}\right)^{(n-1)^\epsilon}\left(1+\frac{(n-1)^\epsilon}{m-1}\right)^{(m-1)-(n-1)^\epsilon} \\
&\leq \left(\frac{1}{(m-1)^{1-\epsilon}}\right)^{\rho^\epsilon(m-1)^\epsilon}\left(1+\frac{1}{(m-1)^{1-\epsilon}}\right)^{m-1} \\
&\leq e^{-(m-1)^\epsilon((1-\epsilon)\rho^\epsilon\log(m-1)-1)} \qquad (6)
\end{aligned}$$

and

$$\begin{aligned}
\left(\left(\frac{(n-2)}{m-1}\right)^{\frac{(n-2)}{m-1}}\left(1+\frac{(n-2)}{m-1}\right)^{1-\frac{(n-2)}{m-1}}\right)^{m-1} &= \left(\frac{(n-2)}{m-1}\right)^{(n-2)}\left(1+\frac{(n-2)}{m-1}\right)^{(m-1)-(n-2)} \\
&\leq \frac{1}{\rho^2}\rho^n(1+\rho)(1+\rho)^{m-n} \\
&\leq \frac{1+\rho}{\rho^2}\left(\rho^\rho(1+\rho)^{1-\rho}\right)^m \\
&\leq \frac{1+\rho}{\rho^2}e^{m(\rho\log\rho+(1-\rho)\log(1+\rho))} \qquad (7)
\end{aligned}$$

Then for m sufficiently large, the case $l=(n-1)^\epsilon$ dominates. Equations 5, 6 and 7 lead to

$$V_{m,m,l} \leq \frac{\sqrt{6}e^2e^{\frac{1}{24}}}{\sqrt{2\pi}}(\rho(m-1))^{-\frac{\epsilon}{2}}e^{-(m-1)^\epsilon(\epsilon\rho^\epsilon\log(m-1)-1)}$$

Thus,

$$\begin{aligned}
\left|\sum_{l=(n-1)^\epsilon}^{n-2} X_{m,n,v,l}\right| &\leq 2((n-2)-(n-1)^\epsilon+1)(m-1)\frac{\sqrt{6}e^2e^{\frac{1}{24}}}{\sqrt{2\pi}}(\rho(m-1))^{-\frac{\epsilon}{2}}e^{-(m-1)^\epsilon(\epsilon\rho^\epsilon\log(m-1)-1)} \\
&\leq 2\frac{\sqrt{6}e^2e^{\frac{1}{24}}}{\sqrt{\pi}}\rho^{1-\frac{\epsilon}{2}}(m-1)^{2-\frac{\epsilon}{2}}e^{-(m-1)^\epsilon(\epsilon\rho^\epsilon\log(m-1)-1)}
\end{aligned}$$

and finally

$$\left|\sum_{l=(n-1)^\epsilon+1}^{(n-2)} X_{m,n,v,l}\right| = o\left(\frac{1}{m^2}\right) \qquad (8)$$

Pooling together equations 1, 2, 3 and 8 gives the following result : $\forall\rho<1$ and $\forall 0<\nu<1$, when $m\to\infty$,

$$P_{m,\rho m,\nu(m-1)} = 1-\frac{4\rho}{m}+O\left(\frac{1}{m^2}\right). \qquad (9)$$

A.2 Case $\rho < 1$ and $\nu = 0$ or 1

Firstly, we can potice that $P_{m,n,v} = P_{m,n,m-1-v}$, or equivalently $P_{m,\rho m,\nu(m-1)} = P_{m,\rho m,(1-\nu)(m-1)}$, by symmetry of the problem. Thus we only need to proove the theorem for $\nu = 0$. Taking $\nu = 0$ leads to

$$P_{m,n,0} = \prod_{l=1}^{n-2} \left(1 - (1 - r_{m,0,l+1,0}) \binom{n-1}{l} p_{m,l+1}^{l} \left((1 - p_{m,l+1})^{n-l-1} - q_{m,l+1,\alpha}^{n-l-1} \right) \right).$$

The sum over α is reduced to $\alpha = 0$, and we have $X_{m,n,0,l} = \log(1 - L_{m,n,0,l,0})$. We cut the sum over l the same way as in the previous case.

– $l = 1, 2, 3$. We have

$$L_{m,\rho m,0,1,0} = \frac{2\rho(1 - e^{-4\rho})}{m} + O\left(\frac{1}{m^2}\right)$$

that is

$$X_{m,\rho m,0,1} = -\frac{2\rho(1 - e^{-4\rho})}{m} + O\left(\frac{1}{m^2}\right). \tag{10}$$

Besides,

$$X_{m,\rho m,0,2} = O\left(\frac{1}{m^2}\right)$$
$$X_{m,\rho m,0,3} = O\left(\frac{1}{m^3}\right) \tag{11}$$

– $3 < l < (n-1)^{\epsilon}$. Reducing the sum over α to $\alpha = 0$, we can apply the result of previous section, that is :

$$\left| \sum_{l=4}^{(n-1)^{\epsilon}-1} X_{m,n,0,l} \right| \leq \frac{4^4 5^4}{12} \frac{\rho^{4+\epsilon}}{m^{4-\epsilon}}$$

and

$$\left| \sum_{l=4}^{(n-1)^{\epsilon}-1} X_{m,n,0,l} \right| = o\left(\frac{1}{m^2}\right) \tag{12}$$

– $(n-1)^{\epsilon} \leq l \leq n-2$. The result of previous section can be used again here. Thus

$$\left| \sum_{l=(n-1)^{\epsilon}}^{(n-2)} X_{m,n,0,l} \right| \leq \frac{3e^2 e^{\frac{1}{24}}}{\sqrt{2\pi}} (\rho(m-1))^{1-\frac{\epsilon}{2}} e^{-(m-1)^{\epsilon}(\epsilon\rho^{\epsilon}\log(m-1)-1)}$$

and

$$\left| \sum_{l=(n-1)^{\epsilon}}^{(n-2)} X_{m,n,0,l} \right| = o\left(\frac{1}{m^2}\right) \tag{13}$$

Pooling together equations 10, 11, 12 and 13 gives the following result : $\forall \rho < 1$, when $m \to \infty$,

$$P_{m,\rho m,0} = P_{m,\rho m,1} = 1 - \frac{2\rho(1 - e^{-4\rho})}{m} + O\left(\frac{1}{m^2}\right). \tag{14}$$

A.3 Case $\rho = 1$ and $\nu = 0$ or 1

Here, we only need to poove the theorem for $\nu = 0$ too. But this time, we have to cut the sum over l this way : $l = 1, 2, 3$; $3 < l < (m-1)^\epsilon$; $(m-1)^\epsilon \leq l \leq (m-1) - (m-1)^\epsilon$, $(m-1) - (m-1)^\epsilon < l \leq m-2$.

- $l = 1, 2, 3$. We have

$$L_{m,m,0,1,0} = \frac{2(1-e^{-4})}{m} + O\left(\frac{1}{m^2}\right).$$

  Then

$$X_{m,m,0,1} = -\frac{2(1-e^{-4})}{m} + O\left(\frac{1}{m^2}\right). \tag{15}$$

  Besides,

$$\begin{aligned} X_{m,m,0,2} &= O\left(\frac{1}{m^2}\right) \\ X_{m,m,0,3} &= O\left(\frac{1}{m^3}\right) \end{aligned} \tag{16}$$

- $3 < l < (m-1)^\epsilon$. Here, we can use the result of previous section, replacing ρ by 1, that is

$$\left|\sum_{l=4}^{(m-1)^\epsilon - 1} X_{m,m,0,l}\right| \leq \frac{4^4 5^4}{12}\frac{1}{m^{4-\epsilon}}.$$

  Thus

$$\left|\sum_{l=4}^{(m-1)^\epsilon - 1} X_{m,m,0,l}\right| = o\left(\frac{1}{m^2}\right). \tag{17}$$

- $(m-1)^\epsilon \leq l \leq (m-1) - (m-1)^\epsilon$. We begin with the following equation (see section A.1) :

$$L_{m,m,0,l,0} \leq \frac{e^2 e^{\frac{1}{24}}}{\sqrt{2\pi}}\frac{1}{\sqrt{m-1}}\frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}}\left(\left(\frac{l}{m-1}\right)^{\frac{l}{m-1}}\left(1+\frac{l}{m-1}\right)^{1-\frac{l}{m-1}}\right)^{m-1}$$

  $\left(\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)\right)^{-1/2}$ is maximal when $l = (m-1)^\epsilon$ and when $l = (m-1) - (m-1)^\epsilon$ (by symmetry). Thus,

$$\begin{aligned} \frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}} &\leq \frac{1}{\sqrt{\frac{1}{(m-1)^{1-\epsilon}}\left(1-\frac{1}{(m-1)^{1-\epsilon}}\right)}} \\ &\leq 2(m-1)^{\frac{1}{2}-\frac{\epsilon}{2}} \end{aligned}$$

  Moreover, we have

$$\begin{aligned} \left(\left(\frac{(m-1)^\epsilon}{m-1}\right)^{\frac{(m-1)^\epsilon}{m-1}}\left(1+\frac{(m-1)^\epsilon}{m-1}\right)^{1-\frac{(m-1)^\epsilon}{m-1}}\right)^{m-1} &= \left(\frac{1}{(m-1)^{1-\epsilon}}\right)^{(m-1)^\epsilon}\left(1+\frac{1}{(m-1)^{1-\epsilon}}\right)^{(m-1)-(m-1)^\epsilon} \\ &\leq e^{-(m-1)^\epsilon((1-\epsilon)\log(m-1)-1)} \end{aligned}$$

and

$$\left(\left(\frac{(m-1)-(m-1)^\epsilon}{m-1}\right)^{\frac{(m-1)-(m-1)^\epsilon}{m-1}}\left(1+\frac{(m-1)-(m-1)^\epsilon}{m-1}\right)^{1-\frac{(m-1)-(m-1)^\epsilon}{m-1}}\right)^{m-1}$$
$$\begin{aligned}&= \left(1-\tfrac{1}{(m-1)^{1-\epsilon}}\right)^{(m-1)-(m-1)^\epsilon}\left(2-\tfrac{1}{(m-1)^{1-\epsilon}}\right)^{(m-1)^\epsilon}\\ &= e^{((m-1)-(m-1)^\epsilon)\log\left(1-\frac{1}{(m-1)^{1-\epsilon}}\right)+(m-1)^\epsilon\log\left(2-\frac{1}{(m-1)^{1-\epsilon}}\right)}\\ &\leq e^{-(m-1)^\epsilon(1-\log 2)+\frac{1}{2(m-1)^{1-2\epsilon}}}\\ &\leq e^{\frac{1}{2}}e^{-(m-1)^\epsilon(1-\log 2)}.\end{aligned}$$

Thus, at least for m sufficiently large,

$$L_{m,m,0,l,0} \leq \frac{2e^2e^{\frac{1}{2}+\frac{1}{24}}}{\sqrt{2\pi}}(m-1)^{-\frac{\epsilon}{2}}e^{-(m-1)^\epsilon(1-\log 2)}$$

and

$$\left|\sum_{l=(m-1)^\epsilon}^{(m-1)-(m-1)^\epsilon-1}\log(1-L_{m,m,0,l,0})\right| \leq \frac{4e^2e^{\frac{1}{2}+\frac{1}{24}}}{\sqrt{2\pi}}(m-1)^{1-\frac{\epsilon}{2}}e^{-(m-1)^\epsilon(1-\log 2)}$$

and finally

$$\left|\sum_{l=(m-1)^\epsilon}^{(m-1)-(m-1)^\epsilon-1}X_{m,m,0,l}\right| = o\left(\frac{1}{m^2}\right). \tag{18}$$

– $(m-1)-(m-1)^\epsilon < l \leq m-2$. The substitution $k = m-1-l$ yields :

$$\sum_{l=(m-1)-(m-1)^\epsilon}^{m-2} X_{m,m,0,l} = \sum_{k=1}^{(m-1)^\epsilon} X_{m,m,0,l}$$

Thus, we have $k = o(\sqrt{m})$, and then

$$L_{m,m,0,m-1-k,0} = \frac{2^ke^{-2k}k^k}{(k-1)!}\left(\frac{1}{m}+O\left(\frac{1}{m^2}\right)\right)$$

Let's set $S_N = \sum\limits_{k=1}^{N}\frac{2^ke^{-2k}k^k}{(k-1)!}$ and $R_N = \sum\limits_{k=N+1}^{\infty}\frac{2^ke^{-2k}k^k}{(k-1)!}$. We have

$$\frac{2^ke^{-2k}k^k}{(k-1)!} \leq \frac{1}{\sqrt{2\pi}}\sqrt{k}\left(\frac{2}{e}\right)^k$$

Thus S_N converges, and we set $S = \sum_{k=1}^{\infty} \frac{2^k e^{-2k} k^k}{(k-1)!} \approx 1.94264$. Moreover,

$$\begin{aligned} R_N &\leq \sum_{k=N+1}^{\infty} \frac{1}{\sqrt{2\pi}} \sqrt{k} \left(\frac{2}{e}\right)^k \\ &\leq \frac{1}{\sqrt{2\pi}} \sum_{k=N+1}^{\infty} k \left(\frac{2}{e}\right)^k \\ &= \frac{1}{\sqrt{2\pi}} \frac{\left(\frac{2}{e}\right)^{N+1} \left(N\left(1-\frac{2}{e}\right)+1\right)}{\left(1-\frac{2}{e}\right)^2} \end{aligned}$$

Therefore $R_{(m-1)^\epsilon} = o\left(\frac{1}{m}\right)$. Finally,

$$\sum_{k=1}^{(m-1)^\epsilon} \log(1 - L_{m,m,0,m-1-k,0}) = -\frac{S}{m} + O\left(\frac{1}{m^2}\right) \tag{19}$$

Pooling together equations 15, 16, 17, 18 and 19 gives the following result : $\forall 0 < \nu < 1$, when $m \to \infty$,

$$P_{m,m,0} = P_{m,m,m-1} = 1 - \frac{2(1-e^{-4}) + S}{m} + O\left(\frac{1}{m^2}\right) \tag{20}$$

A.4 Case $\rho = 1$ and $0 < \nu < 1$

The argumentation is the same as the previous one.

- $l = 1, 2, 3$. As in section A.1, replacing n by m, we have

$$X_{m,m,v,1} = \log(1 - L_{m,m,v,1,v-1}) + \log(1 - L_{m,m,v,1,v})$$

  with

$$L_{m,m,\nu(m-1),1,\nu(m-1)-1} = \frac{2}{m} + O\left(\frac{1}{m^2}\right) \text{ and } L_{m,m,\nu(m-1),1,\nu(m-1)} = \frac{2}{m} + O\left(\frac{1}{m^2}\right)$$

  Thus,

$$X_{m,m,\nu(m-1),1} = -\frac{4}{m} + O\left(\frac{1}{m^2}\right) \tag{21}$$

  Besides, we have

$$X_{m,m,\nu(m-1),2} = O\left(\frac{1}{m^2}\right) \text{ and } X_{m,m,\nu(m-1),2} = O\left(\frac{1}{m^2}\right) \tag{22}$$

- $3 < l < (m-1)^\epsilon$. We can use the results of previous section, with $\rho = 1$, that is

$$\left| \sum_{l=4}^{(m-1)^\epsilon - 1} X_{m,m,v,l} \right| \leq \frac{4^4 5^4}{12} \frac{1}{m^{3-\epsilon}}$$

  Thus,

$$\left| \sum_{l=4}^{(m-1)^\epsilon - 1} X_{m,m,v,l} \right| = o\left(\frac{1}{m^2}\right). \tag{23}$$

- $(m-1)^\epsilon \leq l \leq (m-1)-(m-1)^\epsilon$. We can use the results of previous section, with $\rho = 1$, that is

$$\left|\sum_{l=(m-1)^\epsilon}^{(m-1)-(m-1)^\epsilon-1} X_{m,m,v,l}\right| \leq \frac{4e^2 e^{\frac{1}{2}+\frac{1}{24}}}{\sqrt{2\pi}}(m-1)^{2-\frac{\epsilon}{2}} e^{-(m-1)^\epsilon(1-\log 2)}$$

thus

$$\left|\sum_{l=(m-1)^\epsilon}^{(m-1)-(m-1)^\epsilon-1} X_{m,m,v,l}\right| = o\left(\frac{1}{m^2}\right) \tag{24}$$

- $(m-1)-(m-1)^\epsilon < l \leq m-2$. The substitution $k = m-l-1$ yields

$$\sum_{l=(m-1)-(m-1)^\epsilon}^{m-2} \sum_{\alpha=max(0,v-l)}^{min(v,m-l-1)} \log(1-L_{m,m,v,l,\alpha}) = \sum_{k=1}^{(m-1)^\epsilon} \sum_{\alpha=max(0,v-m+k+1)}^{min(v,k)} \log(1-L_{m,m,v,m-1-k,\alpha})$$

As $k \leq (m-1)^\epsilon$, we have $v \geq k$ $\forall m \geq 1+\nu^{-\frac{1}{1-\epsilon}}$, and $v-m+k+1 \leq 0$ $\forall m \geq 1+(1-\nu)^{-\frac{1}{1-\epsilon}}$. Then

$$\sum_{l=(m-1)-(m-1)^\epsilon}^{m-2} \sum_{\alpha=max(0,v-l)}^{min(v,m-l-1)} \log(1-L_{m,m,v,l,\alpha}) = \sum_{k=1}^{(m-1)^\epsilon} \sum_{\alpha=0}^{k} \log(1-L_{m,m,v,m-1-k,\alpha})$$

Thus we have $k = o(\sqrt{m})$ and $\alpha = o(\sqrt{m})$. Therefore

$$L_{m,m,\nu(m-1),m-1-k,\alpha} = \frac{2^k e^{-2k} k^k (\alpha + k\nu - 2\alpha\nu)}{m\nu(1-\nu)k!} + O\left(\frac{1}{m^2}\right)$$

and

$$\log(L_{m,m,\nu(m-1),m-1-k,\alpha}) = -\frac{2^k e^{-2k} k^k (\alpha + k\nu - 2\alpha\nu)}{m\nu(1-\nu)k!} + O\left(\frac{1}{m^2}\right)$$

Performing the sum over α gives

$$\sum_{\alpha=0}^{k} \frac{2^k e^{-2k} k^k (\alpha + k\nu - 2\alpha\nu)}{m\nu(1-\nu)k!} = \frac{2^k e^{-2k} k^k}{m\nu(1-\nu)k!} \frac{k(k+1)}{2}$$

Let's set $S'_N = \sum_{k=1}^{N} \frac{2^k e^{-2k} k^k (k+1)}{(k-1)!}$ and $R'_N = \sum_{k=N+1}^{\infty} \frac{2^k e^{-2k} k^k (k+1)}{(k-1)!}$. We have

$$\frac{2^k e^{-2k} k^k (k+1)}{(k-1)!} \leq \frac{1}{\sqrt{2\pi}} \sqrt{k}(k+1) \left(\frac{2}{e}\right)^k$$

Thus S'_N converges, and we set $S' = \sum_{k=1}^{\infty} \frac{2^k e^{-2k} k^k (k+1)}{(k-1)!} \approx 11.9359$. Moreover,

$$\begin{aligned} R'_N &\leq \sum_{k=N+1}^{\infty} \frac{1}{\sqrt{2\pi}} \sqrt{k}(k+1) \left(\frac{2}{e}\right)^k \\ &\leq \frac{1}{\sqrt{2\pi}} \sum_{k=N+1}^{\infty} k(k+1) \left(\frac{2}{e}\right)^k \\ &= \frac{1}{\sqrt{2\pi}} \frac{\left(\frac{2}{e}\right)^{N+1} Q(N)}{\left(1-\frac{2}{e}\right)^3} \end{aligned}$$

where $Q(N)$ is a polynom of degree 2 in N. Then $R'_{(m-1)^\epsilon} = o\left(\frac{1}{m}\right)$. To conclude,

$$\sum_{l=(m-1)-(m-1)^\epsilon}^{m-2} X_{m,m,\nu(m-1),l} = -\frac{S'}{2\nu(1-\nu)m} + O\left(\frac{1}{m^2}\right) \tag{25}$$

Finally, pooling together equations 21, 22, 23, 24 and 25 gives the following result : $\forall 0 < \nu < 1$, when $m \to \infty$,

$$\boxed{P_{m,m,\nu(m-1)} = 1 - \frac{4 + \frac{S'}{2\nu(1-\nu)}}{m} + O\left(\frac{1}{m^2}\right)} \tag{26}$$

B Proof of theorem 2

We set $D_{m,n,v,l,\alpha}(d) = (1 - r_{m,v,l+1,\alpha}(d))\binom{n-1}{l}p^l_{m,l+1}(d)\left((1-p_{m,l+1}(d))^{n-l-1} - q^{n-l-1}_{m,l,\alpha}(d)\right)$. Notice that for $l < d-1$, $p_{m,l+1}(d) = 0$. Thus,

$$\begin{aligned} P_{m,n,v}(d) &= \prod_{l=d-1}^{n-2} \prod_{\alpha=max(0,v-l)}^{min(v,m-l-1)} (1 - D_{m,n,v,l,\alpha}(d)) \\ &= \exp\left(\sum_{l=d-1}^{n-2} \sum_{\alpha=max(0,v-l)}^{min(v,m-l-1)} \log(1 - D_{m,n,v,l,\alpha}(d))\right) \end{aligned}$$

We also define $X_{m,n,v,l}(d) = \sum\limits_{\alpha=max(0,v-l)}^{min(v,m-l-1)} \log(1 - D_{m,n,v,l,\alpha}(d))$, such that $P_{m,n,v}(d) = \exp\left(\sum\limits_{l=d-1}^{n-2} X_{m,n,v,l}(d)\right)$. Let's set $\rho = n/m$, $\nu = v/(m-1)$ and $\mu = d/m$. We have $0 < \rho \leq 1$, $0 \leq \nu \leq 1$ and $0 < \mu \leq 1$. We distinguish the 3 cases : $\nu = 0$ or $\nu = 1$ or $\mu \geq \rho$; $0 < \rho < 1$ and $0 < \nu < 1$; $\rho = 1$ and $0 < \nu < 1$. We may use some results obtained in the proof of theorem 2.

B.1 $\nu = 0$ or $\nu = 1$ or $\mu \geq \rho$

As $P_{m,n,v}(d) = P_{m,n,m-1-v}(d)$, we only need to proove the result for $v = 0$, that is $\nu = 0$. We have

$$P_{m,n,0}(d) = \prod_{l=d-1}^{n-2}\left(1 - (1 - r_{m,0,l+1,0}(d))\binom{n-1}{l}p^l_{m,l+1}(d)\left((1-p_{m,l+1})^{n-l-1} - q^{n-l-1}_{m,l+1,0}\right)\right)$$

with $r_{m,0,l+1,0}(d) = \frac{min(0,l+1-d)-max(0,-d+1)+1}{min(0,m-d)-max(0,-d+1)+1}$. We have $d \geq 1$, $m \geq d$ and $l+1 > d$, thus all the min and max equal 0, that is $r_{m,0,l+1,0}(d) = 1$. Therefore, $\forall 0 < \rho \leq 1$ et $\forall 0 < \mu \leq 1$,

$$P_{m,\rho m,0}(\mu m) = P_{m,\rho m,m-1}(\mu m) = 1 \tag{27}$$

The case $\mu \geq \rho$ is trivial : the product over l goes from $d-1$ to $n-2$. Then, if $d \geq n$, the product is reduced to 1. Thus, $\forall \mu \geq \rho$,

$$P_{m,\rho m,\nu(m-1)}(\mu m) = 1 \tag{28}$$

B.2 $0 < \rho < 1$ and $0 < \nu < 1$

As $r_{m,v,l+1,\alpha} \leq 1$ and $q_{m,l+1,\alpha} \leq 1 - p_{l+1,m}$, we have

$$D_{m,n,v,l,\alpha}(d) \leq W_{m,n,l}(d) := \binom{n-1}{l} p_{m,l+1}^{l}(d) \left(1 - p_{m,l+1}(d)\right)^{n-l-1}.$$

We have

$$\begin{aligned}
\frac{W_{m,n+1,l}(d)}{W_{m,n,l}(d)} &= \frac{n}{n-l}\left(1 - \frac{l-d+2}{m-d+1}\right) \\
&= \frac{n}{n-l}\left(\frac{m-l-1}{m-d+1}\right)
\end{aligned}$$

We can see that $\frac{W_{m,n+1,l}(d)}{W_{m,n,l}(d)}$ decreases with l. Therefore,

$$\begin{aligned}
\frac{W_{m,n+1,l}(d)}{W_{m,n,l}(d)} \geq \frac{W_{m,n+1,d-1}(d)}{W_{m,n,d-1}(d)} &= \frac{n}{n-d+1}\frac{m-d}{m-d+1} \\
&= 1 + \frac{(d-1)(m-d+1) - (n-d+1)}{(n-d+1)(m-d+1)} \\
&\geq 1 + \frac{m-n}{(n-d+1)(m-d+1)} \qquad \text{because } d-1 \geq 1 \\
&\geq 1
\end{aligned}$$

Thus, W increases with n, and is bounded by $W_{m,m,l}(d)$. Then we have

$$D_{m,n,v,l,\alpha}(d) \leq W_{m,n,l}(d) = \binom{m-1}{l} p_{m,l+1}^{l}(d) \left(1 - p_{m,l+1}(d)\right)^{m-l-1}$$

One the one hand, we have

$$\binom{m-1}{l} \leq \frac{e^{\frac{1}{24}}}{\sqrt{2\pi}} \frac{1}{\sqrt{m-1}} \frac{1}{\sqrt{\frac{l}{m-1}\left(1 - \frac{l}{m-1}\right)}} \frac{1}{\left(\frac{l}{m-1}\right)^{l}\left(1 - \frac{l}{m-1}\right)^{m-l-1}}$$

On the other hand,

$$\begin{aligned}
\left(\frac{l-d+2}{m-d+1}\right)^{l} &= \left(\frac{(m-1)\frac{l}{m-1} - m\frac{d-2}{m}}{m - m\frac{d-1}{m}}\right)^{l} \\
&\leq \left(\frac{\frac{l}{m-1} - \frac{d-2}{m}}{1 - \frac{d}{m}}\right)^{l}
\end{aligned}$$

and

$$
\begin{aligned}
\left(1-\frac{l-d+2}{m-d+1}\right)^{m-l-1} &= \left(\frac{m-l-1}{m-d+1}\right)^{m-l-1} \\
&= \left(\frac{(m-1)-(m-1)\frac{l}{m-1}}{m-m\frac{d-1}{m}}\right)^{m-l-1} \\
&\leq \left(\frac{m-m\frac{l}{m-1}}{m-m\frac{d}{m}}\right)^{m-l-1} \\
&\leq \left(\frac{1-\frac{l}{m-1}}{1-\frac{d}{m}}\right)^{m-l-1}
\end{aligned}
$$

Thus,

$$
\begin{aligned}
D_{m,n,v,l,\alpha}(d) &\leq \frac{e^{\frac{1}{24}}}{\sqrt{2\pi}}\frac{1}{\sqrt{m-1}}\frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}}\frac{\left(\frac{\frac{l}{m-1}-\frac{d-2}{m}}{1-\frac{d}{m}}\right)^{l}\left(\frac{1-\frac{l}{m-1}}{1-\frac{d}{m}}\right)^{m-l-1}}{\left(\frac{l}{m-1}\right)^{l}\left(1-\frac{l}{m-1}\right)^{m-l-1}} \\
&\leq \frac{e^{\frac{1}{24}}}{\sqrt{2\pi}}\frac{1}{\sqrt{m-1}}\frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}}\left(1-\frac{\frac{d-2}{m}}{\frac{l}{m-1}}\right)^{l}\left(\frac{1}{1-\mu}\right)^{m-1}
\end{aligned}
$$

We can see that $\left(1-\frac{\frac{d-2}{m}}{\frac{l}{m-1}}\right)^{l}$ increases with l. Thus,

$$
\begin{aligned}
\left(1-\frac{\frac{d-2}{m}}{\frac{l}{m-1}}\right)^{l} &\leq \left(1-\frac{\frac{d-2}{m}}{\frac{n-2}{m-1}}\right)^{n-2} \\
&= \left(\frac{m(n-d)+(d-2)}{m(n-2)}\right)^{n-2} \\
&\leq \left(1-\frac{\mu}{\rho}\right)^{n-2}\left(1+\frac{2}{n-2}\right)^{n-2}\left(1+\frac{d-2}{m(n-d)}\right)^{n-2} \\
&\leq \left(1-\frac{\mu}{\rho}\right)^{n-2} e^{2} e^{\frac{\rho\mu}{\rho-\mu}} \\
&\leq \left(\frac{\rho}{\rho-\mu}\right)^{2}\left(1-\frac{\mu}{\rho}\right)^{\rho(m-1)} e^{2} e^{\frac{\rho\mu}{\rho-\mu}}
\end{aligned}
$$

Moreover, $\left(\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)\right)^{-1/2}$ is maximal when $l = d-1$ or $l = n-2$, with

$$\frac{1}{\sqrt{\frac{d-1}{m-1}\left(1-\frac{d-1}{m-1}\right)}} \leq \frac{\sqrt{2}}{\sqrt{\mu(1-\mu)}}$$
$$\frac{1}{\sqrt{\frac{n-2}{m-1}\left(1-\frac{n-2}{m-1}\right)}} \leq \frac{\sqrt{3}}{\sqrt{\rho(1-\rho)}}$$

Let's set $C_{\mu,\rho} = max\left(\frac{\sqrt{2}}{\sqrt{\mu(1-\mu)}}, \frac{\sqrt{3}}{\sqrt{\rho(1-\rho)}}\right)$. then,

$$D_{m,n,v,l,\alpha}(d) \leq \frac{e^{2+\frac{1}{24}}}{\sqrt{2\pi}}\left(\frac{\rho}{\rho-\mu}\right)^2 e^{\frac{\rho\mu}{\rho-\mu}} C_{\mu,\rho}\frac{1}{\sqrt{m-1}} e^{-(m-1)\left(\rho\log\left(\frac{\rho}{\rho-\mu}\right)+\log(1-\mu)\right)}$$

Thus,

$$\left|\sum_{l=d-1}^{n-2} X_{m,n,v,l}\right| \leq 2\sum_{l=d-1}^{n-2}\sum_{\alpha=max(0,v-l)}^{min(v,m-l-1)} D_{m,n,v,l,\alpha}(d)$$
$$\leq 2\frac{e^{2+\frac{1}{24}}}{\sqrt{2\pi}}\frac{\rho^2}{\rho-\mu} e^{\frac{\rho\mu}{\rho-\mu}} C_{\mu,\rho}(m-1)^{\frac{3}{2}} e^{-(m-1)\left(\rho\log\left(\frac{\rho}{\rho-\mu}\right)+\log(1-\mu)\right)}$$

and finally, $\forall 0 < \mu < \rho < 1$,

$$P_{m,\rho m,\nu(m-1)}(\mu m) \geq 1 - 2\frac{e^{2+\frac{1}{24}}}{\sqrt{2\pi}}\frac{\rho^2}{\rho-\mu} e^{\frac{\rho\mu}{\rho-\mu}} C_{\mu,\rho}(m-1)^{\frac{3}{2}} e^{-(m-1)\left(\rho\log\left(\frac{\rho}{\rho-\mu}\right)+\log(1-\mu)\right)} \quad (29)$$

in particular

$$P_{m,\rho m,\nu(m-1)}(\mu m) = o\left(\frac{1}{m^2}\right) \quad (30)$$

B.3 $\rho = 1$ and $0 < \nu < 1$

Let $0 < \epsilon < 1/2$. We cut the sum over l as follows : $d-1 \leq l < (m-1)-(m-1)^\epsilon$ and $(m-1)-(m-1)^\epsilon \leq l \leq m-2$.

- $d-1 \leq l < (m-1)-(m-1)^\epsilon$. From the previous section, we have

$$D_{m,m,v,l,\alpha}(d) \leq \frac{e^{\frac{1}{24}}}{\sqrt{2\pi}}\frac{1}{\sqrt{m-1}}\frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}}\left(1-\frac{\frac{d-2}{m}}{\frac{l}{m-1}}\right)^l\left(\frac{1}{1-\mu}\right)^{m-1}$$

$\left(1-\frac{\frac{d-2}{m}}{\frac{l}{m-1}}\right)^l$ increases with l. Thus, $l < (m-1)-(m-1)^\epsilon)$ implies

$$\begin{aligned}
\left(1-\frac{\frac{d-2}{m}}{\frac{l}{m-1}}\right)^l &\leq \left(1-\frac{\frac{d-2}{m}}{\frac{(m-1)-(m-1)^\epsilon}{m-1}}\right)^{(m-1)-(m-1)^\epsilon} \\
&= e^{((m-1)-(m-1)^\epsilon)\log\left(1-\frac{(m-1)(d-2)}{m((m-1)-(m-1)^\epsilon)}\right)} \\
&= e^{((m-1)-(m-1)^\epsilon)\log\left(1-\mu\left(1-\frac{2}{\mu m}\right)\frac{(m-1)}{m((m-1)-(m-1)^\epsilon)}\right)} \\
&= e^{((m-1)-(m-1)^\epsilon)\log\left((1-\mu)\left(1-\frac{1}{1-\mu}\left(\frac{(m-1)^\epsilon}{(m-1)-(m-1)^\epsilon}-\frac{2}{\mu m}\left(1-\frac{(m-1)^\epsilon}{(m-1)-(m-1)^\epsilon}\right)\right)\right)\right)} \\
&\leq e^{((m-1)-(m-1)^\epsilon)\log(1-\mu))-\frac{(m-1)^\epsilon}{1-\mu}+\frac{2}{\mu}\left(1+\frac{1}{(m-1)^{1-\epsilon}}\right)} \\
&\leq e^{\frac{2+\sqrt{2}}{\mu}} e^{((m-1)-(m-1)^\epsilon)\log(1-\mu))-\frac{(m-1)^\epsilon}{1-\mu}}
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
\frac{1}{\sqrt{\frac{l}{m-1}\left(1-\frac{l}{m-1}\right)}} &\leq \frac{1}{\sqrt{\frac{(m-1)-(m-1)^\epsilon}{m-1}\left(1-\frac{(m-1)-(m-1)^\epsilon}{m-1}\right)}} \\
&= \frac{1}{\sqrt{\left(1-\frac{1}{(m-1)^{1-\epsilon}}\right)\frac{1}{(m-1)^{1-\epsilon}}}} \\
&\leq \sqrt{2+\sqrt{2}}(m-1)^{\frac{1}{2}-\frac{\epsilon}{2}}
\end{aligned}$$

Thus,

$$D_{m,m,v,l,\alpha}(d) \leq \frac{\sqrt{2+\sqrt{2}}e^{\frac{1}{24}}}{\sqrt{2\pi}} e^{\frac{2+\sqrt{2}}{\mu}}(m-1)^{-\frac{\epsilon}{2}} e^{-(m-1)^\epsilon\left(\log(1-\mu)+\frac{1}{1-\mu}\right)}.$$

Therefore

$$\left|\sum_{l=d-1}^{(m-1)-(m-1)^\epsilon} X_{m,m,v,l}(d)\right| \leq \frac{2\sqrt{2+\sqrt{2}}e^{\frac{1}{24}}}{\sqrt{2\pi}} e^{\frac{2+\sqrt{2}}{\mu}}(m-1)^{2-\frac{\epsilon}{2}} e^{-(m-1)^\epsilon\left(\log(1-\mu)+\frac{1}{1-\mu}\right)}$$

and finally

$$\left|\sum_{l=d-1}^{(m-1)-(m-1)^\epsilon} X_{m,m,v,l}(d)\right| = o\left(\frac{1}{m^2}\right) \tag{31}$$

– $(m-1)-(m-1)^\epsilon \leq l \leq m-2$. The substitution $k = m-1-l$ yields, for m sufficiently large,

$$\sum_{l=(m-1)-(m-1)^\epsilon}^{m-2} \sum_{\alpha=max(0,v-l)}^{min(v,m-l-1)} \log(1-D_{m,m,v,l,\alpha}(d)) = \sum_{k=1}^{(m-1)^\epsilon}\sum_{\alpha=0}^{k}\log(1-D_{m,m,v,m-1-k,\alpha}(d))$$

We have

$$q_{m,m-k,\alpha}(d) = max\left(0, \frac{\alpha-d+1}{m-d+1}\right) + max\left(0, \frac{k-\alpha-d-1}{m-d+1}\right)$$

As $\alpha \leq k \leq (m-1)^\epsilon$ and $k-\alpha \leq k \leq (m-1)^\epsilon$, there exists m_ϵ such that $\forall m \geq m_\epsilon$, $q_{m,m-k,\alpha}(\mu m) = 0$. In addition, we have

$$r_{m,v,m-k,\alpha}(d) = \frac{min(v, m+\alpha-k-d) - max(\alpha, v-d+1)+1}{min(v, m-d) - max(0, v-d+1)+1}$$

The denominator is written in terms of ν and μ as $min(\nu(m-1), m-\mu m) - max(0, \nu(m-1) - \mu m + 1) + 1$. We can see that for m sufficiently large,

$$min(\nu(m-1), m-\mu m) - max(0, \nu(m-1)-\mu m+1)+1 = m\Phi_{\nu\mu} + \chi_{\nu\mu}$$

where $\Phi_{\nu\mu} = min(\nu, 1-\mu) - max(0, \nu-\mu)$, and

$$\chi_{\nu\mu} = \begin{cases} 1 & \text{if } \nu < \mu \text{ and } \nu > 1-\mu \\ \nu & \text{if } \nu \geq \mu \text{ and } \nu > 1-\mu \\ 0 & \text{if } \nu \geq \mu \text{ and } \nu \leq 1-\mu \\ 1-\nu & \text{if } \nu < \mu \text{ and } \nu \leq 1-\mu \end{cases}$$

In the same way, as $\alpha \leq k \leq (m-1)^\epsilon$ and $k-\alpha \leq k \leq (m-1)^\epsilon$, there exists m sufficiently large suche that

$$\begin{aligned} &min(\nu(m-1), m+\alpha-k-\mu m) - max(\alpha, \nu(m-1)-\mu m+1)+1 \\ &\quad = m\Phi_{\nu\mu} + \chi_{\nu\mu} - \xi(\nu,\mu,\alpha,k) - \delta_{\nu,1-\mu}\psi(\nu, k-\alpha) - \delta_{\nu,\mu}\psi(1-\nu,\alpha) \end{aligned}$$

where

$$\xi(\nu,\mu,\alpha,k) = \begin{cases} k & \text{if } \nu < \mu \text{ and } \nu > 1-\mu \\ k-\alpha & \text{if } \nu \geq \mu \text{ and } \nu > 1-\mu \\ 0 & \text{if } \nu \geq \mu \text{ and } \nu \leq 1-\mu \\ \alpha & \text{if } \nu < \mu \text{ and } \nu \leq 1-\mu \end{cases}$$

and $\psi(x,y) = max(0, y-x)$, that is $\psi(\nu, k-\alpha) = max(0, k-\alpha-\nu)$ and $\psi(1-\nu,\alpha) = max(0, \alpha+\nu-1)$. Thus,

$$r_{m,\nu(m-1),m-k,\alpha}(\mu m) = \frac{m\Phi_{\nu\mu} + \chi_{\nu\mu} - \xi(\nu,\mu,\alpha,k) - \delta_{\nu,1-\mu}\psi(\nu,k-\alpha) - \delta_{\nu,\mu}\psi(1-\nu,\alpha)}{m\Phi_{\nu\mu} + \chi_{\nu\mu}}$$

then

$$1 - r_{m,\nu(m-1),m-k,\alpha}(\mu m) = \frac{\xi(\nu,\mu,\alpha,k) + \delta_{\nu,1-\mu}\psi(\nu,k-\alpha) + \delta_{\nu,\mu}\psi(1-\nu,\alpha)}{m\Phi_{\nu\mu}} \left(1 + O\left(\frac{1}{m}\right)\right).$$

In addition, we have

$$\binom{m-1}{k} \left(\frac{m-k-d+1}{m-d+1}\right)^{m-k-1} \left(1 - \frac{m-k-d+1}{m-d+1}\right)^k = \left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{k!} \left(1 + O\left(\frac{1}{m}\right)\right).$$

It yields

$$D_{m,m,\nu(m-1),m-k-1,\alpha}(\mu m) = \frac{\phi(\nu,\mu,\alpha,k) + \delta_{\nu,1-\mu}\psi(\nu,k-\alpha) + \delta_{\nu,\mu}\psi(1-\nu,\alpha)}{m\Xi(\nu,\mu)} \left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{k!} \left(1 + O\left(\frac{1}{m}\right)\right).$$

Let's set $\Xi(\nu,\mu,k) = \sum_{\alpha=0}^{k} \xi(\nu,\mu,\alpha,k)$. We have

$$\Xi(\nu,\mu,k) = \begin{cases} k(k+1) & \text{if } \nu < \mu \text{ and } \nu > 1-\mu \\ \frac{k(k+1)}{2} & \text{if } \nu \geq \mu \text{ and } \nu > 1-\mu \\ 0 & \text{if } \nu \geq \mu \text{ and } \nu \leq 1-\mu \\ \frac{k(k+1)}{2} & \text{if } \nu < \mu \text{ and } \nu \leq 1-\mu \end{cases}$$

Let's set $\Psi(\nu,k) = \sum\limits_{\alpha=0}^{k} \psi(\nu, k-\alpha)$. We have $\Psi(\nu,k) = \frac{k(k+1)}{2} - k\nu$. We also have $\sum\limits_{\alpha=0}^{k} \psi(1-\nu,\alpha) = \Psi(1-\nu,k)$. Then we set $f_N(\mu) = \sum\limits_{k=1}^{N}(k+1)\left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{(k-1)!}$ and $g_N(\mu) = \sum\limits_{k=1}^{N}\left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{(k-1)!}$. Thoses 2 series converge $\forall 0<\mu<1$, and their remainder term equal $o\left(\frac{1}{m}\right)$. We denote the limits

$$f(\mu) = \sum_{k=1}^{\infty}(k+1)\left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{(k-1)!} \text{ and } g(\mu) = \sum_{k=1}^{\infty}\left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{(k-1)!}.$$

Finally, we set $T_{\nu\mu} = \sum\limits_{k=1}^{\infty}\left(\xi(\nu,\mu,k) + \delta_{\nu,1-\mu}\psi(\nu,k) + \delta_{\nu,\mu}\psi(1-\nu,k)\right)\left(\frac{k}{1-\mu}\right)^k \frac{e^{-\frac{k}{1-\mu}}}{k!}$. Then,

$$\begin{aligned}
T_{\nu\mu} &= \begin{cases}
f(\mu) & \text{if } \nu<\mu \text{ and } \nu>1-\mu \\
\frac{f(\mu)}{2} + \delta_{\nu,\mu}\left(\frac{f(\mu)}{2} - (1-\nu)g(\mu)\right) & \text{if } \nu\geq\mu \text{ and } \nu>1-\mu \\
\delta_{\nu,\mu}\left(\frac{f(\mu)}{2} - (1-\nu)g(\mu)\right) + \delta_{\nu,1-\mu}\left(\frac{f(\mu)}{2} - \nu g(\mu)\right) & \text{if } \nu\geq\mu \text{ and } \nu\leq 1-\mu \\
\frac{f(\mu)}{2} + \delta_{\nu,1-\mu}\left(\frac{f(\mu)}{2} - \nu g(\mu)\right) & \text{if } \nu<\mu \text{ and } \nu\leq 1-\mu
\end{cases} \\
&= \begin{cases}
f(\mu) & \text{if } \nu<\mu \text{ and } \nu>1-\mu \\
\frac{f(\mu)}{2} + \delta_{\nu,\mu}\left(\frac{f(\mu)}{2} - (1-\mu)g(\mu)\right) & \text{if } \nu\geq\mu \text{ and } \nu>1-\mu \\
(\delta_{\nu,\mu} + \delta_{\nu,1-\mu})\left(\frac{f(\mu)}{2} - (1-\mu)g(\mu)\right) & \text{if } \nu\geq\mu \text{ and } \nu\leq 1-\mu \\
\frac{f(\mu)}{2} + \delta_{\nu,1-\mu}\left(\frac{f(\mu)}{2} - (1-\mu)g(\mu)\right) & \text{if } \nu<\mu \text{ and } \nu\leq 1-\mu
\end{cases}
\end{aligned}$$

and we have

$$\sum_{l=(m-1)-(m-1)^{\epsilon}}^{m-2} X_{m,m,\nu(m-1),l}(\mu m) = -\frac{T_{\nu\mu}}{m\Phi_{\nu\mu}} + O\left(\frac{1}{m^2}\right). \tag{32}$$

Therefore, equations 31 and 32 give : $\forall 0<\nu<1$ and $\forall \mu<\rho$,

$$P_{m,m,\nu(m-1)}(\mu m) = 1 - \frac{T_{\nu\mu}}{m\Phi_{\nu\mu}} + O\left(\frac{1}{m^2}\right). \tag{33}$$